PROBABILIDADE E ESTATÍSTICA MATEMÁTICA I

ARA, A.

ara@ufpr.br

Slide 02

DISTRIBUIÇÕES CONJUNTAS

DISTRIBUIÇÕES CONTÍNUAS

Anteriormente foram vistas algumas distribuições de probabilidade para variáveis aleatórias discretas e contínuas para apenas uma variável.

- Contexto univariado;

Nesta aula, vamos introduzir a teoria de probabilidades na presença de mais de uma variável.

- Contexto multivariado.

DISTRIBUIÇÕES CONJUNTAS

Distribuição acumulada conjunta

Definição: Sejam $X_1, X_2, \ldots, X_k$ k variáveis aleatórias definidas no mesmo espaço de probabilidade. A função de distribuição acumulada conjunta de $X_1, X_2, \ldots, X_k$, denotada por $F_{X_1,X_2,\ldots,X_k}(\cdot, \cdot, \ldots, \cdot)$, é definida como $P[X_1 \leq x_1, X_2 \leq x_2, \ldots, X_k \leq x_k]$ para todo $(x_1, x_2, \ldots, x_k)$.

- Note que uma função de distribuição acumulada conjunta é uma função com domínio do espaço euclidiano k-dimensional e contradomínio do intervalo $[0, 1]$.

DISTRIBUIÇÕES CONJUNTAS

Considere o experimento de lançarmos dois dados tradicionais de seis faces, como um dado branco e um dado preto, o espaço amostral será:

$$\left\{\begin{array}{cccccc} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{array}\right\}$$

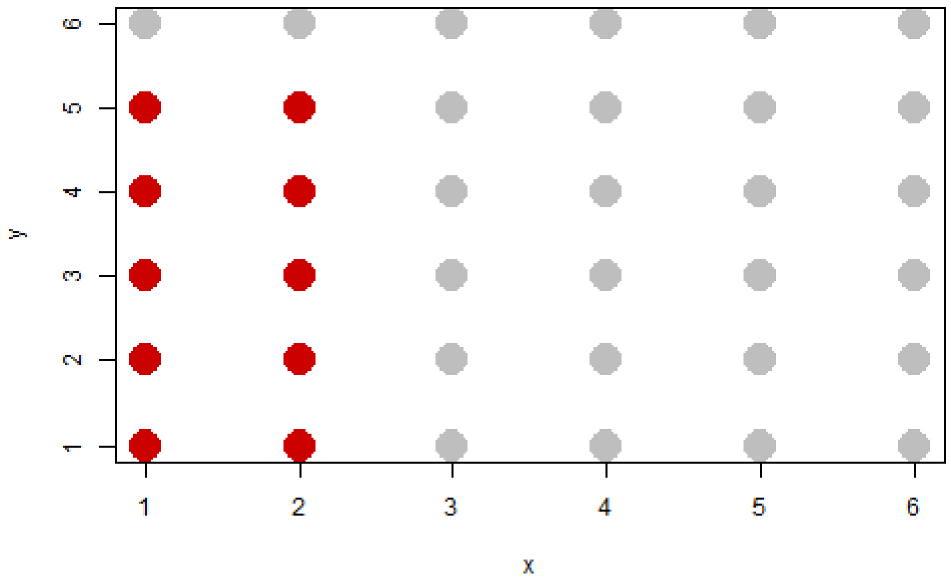
Os 36 pontos amostrais são igualmente prováveis.

Sejam as variáveis aleatórias X=resultado do dado branco e Y= resultado do dado preto.

DISTRIBUIÇÕES CONJUNTAS

Seja $F_{X,Y}(x,y) = P(X \leq x, Y \leq y)$

Poderíamos estar interessados na $P[X \leq 2, Y \leq 5] = F_{X,Y}(2,5)$



DISTRIBUIÇÕES CONJUNTAS

Tabela dos valores de $F_{X,Y}(x,y)$

$y \backslash x$	$x < 1$	$1 \leq x < 2$	$2 \leq x < 3$	$3 \leq x < 4$	$4 \leq x < 5$	$5 \leq x < 6$	$x \leq 6$
$y < 1$	0	0	0	0	0	0	0
$1 \leq y < 2$	0	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$
$2 \leq y < 3$	0	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{6}{36}$	$\frac{8}{36}$	$\frac{10}{36}$	$\frac{12}{36}$
$3 \leq y < 4$	0	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{9}{36}$	$\frac{12}{36}$	$\frac{15}{36}$	$\frac{18}{36}$
$4 \leq y < 5$	0	$\frac{4}{36}$	$\frac{8}{36}$	$\frac{12}{36}$	$\frac{16}{36}$	$\frac{20}{36}$	$\frac{24}{36}$
$5 \leq y < 6$	0	$\frac{5}{36}$	$\frac{10}{36}$	$\frac{15}{36}$	$\frac{20}{36}$	$\frac{25}{36}$	$\frac{30}{36}$
$y \leq 6$	0	$\frac{6}{36}$	$\frac{12}{36}$	$\frac{18}{36}$	$\frac{24}{36}$	$\frac{30}{36}$	1

DISTRIBUIÇÕES CONJUNTAS

Propriedades da função de distribuição cumulativa bivariada $F(\cdot,\cdot)$:

i.

$$F(-\infty, y) = \lim_{x \to -\infty} F(x, y) = 0, \ \forall \ y$$

$$F(x, -\infty) = \lim_{y \to -\infty} F(x, y) = 0, \ \forall \ x$$

$$F(\infty, \infty) = \lim_{\substack{x \to \infty \\ y \to \infty}} F(x, y) = 1, \ \forall \ x$$

ii. Se $x_1 < x_2$ e $y_1 < y_2$, então

$P[x_1 < X \leq x_2, y_1 < Y \leq y_2] = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) > 0$.

DISTRIBUIÇÕES CONJUNTAS

Propriedades da função de distribuição cumulativa bivariada $F(\cdot,\cdot)$:

iii. $F(x, y)$ é contínua à direita, ou seja

$$\lim_{0<h\to 0} F(x+h, y) = \lim_{0<h\to 0} F(x, y+h) = F(x, y)$$

Função de distribuição acumulada bivariada: Qualquer função que satisfaça as propriedades de i a iii é definida como uma função de distribuição acumulada bivariada.

DISTRIBUIÇÕES CONJUNTAS

Distribuições marginais:

O conhecimento sobre a função de distribuição acumulada conjunta de X e Y implica no conhecimento das duas funções de distribuição acumulada marginal.

$$F_X(x) = F_{X,Y}(x, \infty)$$

$$F_Y(y) = F_{X,Y}(\infty, y)$$

Função probabilidade conjunta para variáveis aleatórias discretas

Definição: $(X_1, \ldots, X_k)$ será uma variável aleatória discreta k-dimensional (vetor aleatório) se seus possíveis valores forem finitos ou infinitos enumeráveis.

Isto é, os valores possíveis de $(X_1, \ldots, X_k)$, denotados por $(x_1, \ldots, x_k)$, são finitos ou infinitos enumeráveis.

Definição: Seja $(X_1, \ldots, X_k)$ uma variável aleatória discreta k-dimensional, a cada resultado possível $(x_1, \ldots, x_k)$ associaremos

$$p(x_1, x_2, \ldots, x_k) = P(X_1 = x_1, X2 = x_2, \ldots, X_k = x_k)$$

Com as seguintes propriedades:

i. $p(x_1, x_2, \ldots, x_k) \geq 0, \quad \forall\, x_i, i = 1, \ldots, k$

ii. $\sum p(x_1, x_2, \ldots, x_k) = 1$, sendo a soma sobre todos valores possíveis de $(X_1, \ldots, X_k)$.

Distribuição Multinomial

Distribuição Multinominal

Considere um experimento com k possíveis resultados, o i-ésimo resultado com probabilidade θ_i, $\sum_{i=1}^{k} \theta_i = 1$. X_i representa o número de ocorrências do i-ésimo resultado em n repetições independentes.

$$P(X_1 = x_1, \ldots, X_k = x_k | n, \theta_1, \ldots, \theta_k) = \frac{n!}{x_1! x_2! \ldots x_k!} \theta_1^{x_1} \theta_2^{x_2} \ldots \theta_k^{x_k}$$

em que $\sum_{i=1}^{k} x_i = n$ e $\sum_{i=1}^{k} \theta_i = 1$.

Quando $k = 2$ temos o caso da distribuição binomial.

Função probabilidade conjunta para variáveis aleatórias contínuas

Definição: $(X_1, \ldots, X_k)$ será uma variável aleatória contínua k-dimensional (vetor aleatório) se seus possíveis valores estão em algum conjunto não-enumerável do plano euclidiano.

Por exemplo, se $(X_1, \ldots, X_k)$ puder tomar todos os valores no hiperretângulo $\{(x_1, \ldots, x_k) \mid a_1 \leq x_1 \leq b_1, \ldots,\ a_k \leq x_k \leq b_k\}$.

Isto é, os valores possíveis de $(X_1, \ldots, X_k)$, denotados por $(x_1, \ldots, x_k)$ são finitos ou infinitos enumeráveis.

Função probabilidade conjunta para variáveis aleatórias contínuas

Definição: Seja $(X_1, \ldots, X_k)$ uma variável aleatória contínua k-dimensional, a cada resultado possível $(x_1, \ldots, x_k)$ associaremos uma função densidade de probabilidade k-dimensional $f(x_1, x_2, \ldots, x_k)$

com as seguintes propriedades

i. $f(x_1, x_2, \ldots, x_k) \geq 0, \quad \forall\, x_i, i = 1, \ldots, k$

ii. $\int \ldots \int f(x_1, x_2, \ldots, x_k) dx_1 \ldots dx_k = 1$

Dessa forma,

$$P(a_1 \leq x_1 \leq b_1, \ldots,\ a_k \leq x_k \leq b_k) = \int_{a_1}^{b_1} \ldots \int_{a_k}^{b_k} f(x_1, x_2, \ldots, x_k) dx_1 \ldots dx_k$$

EXEMPLO: Considere a função bivariada dada por

$$f(x,y) = c(x+y)I_{(0,1)}(x)I_{(0,1)}(y) = c(x+y)/I_U(x,y)$$

sendo $U = \{(x,y)|0 < x < 1 \text{ e } 0 < y < 1\}$. Calcule o valor de c para $f(x,y)$ ser considerada uma densidade.

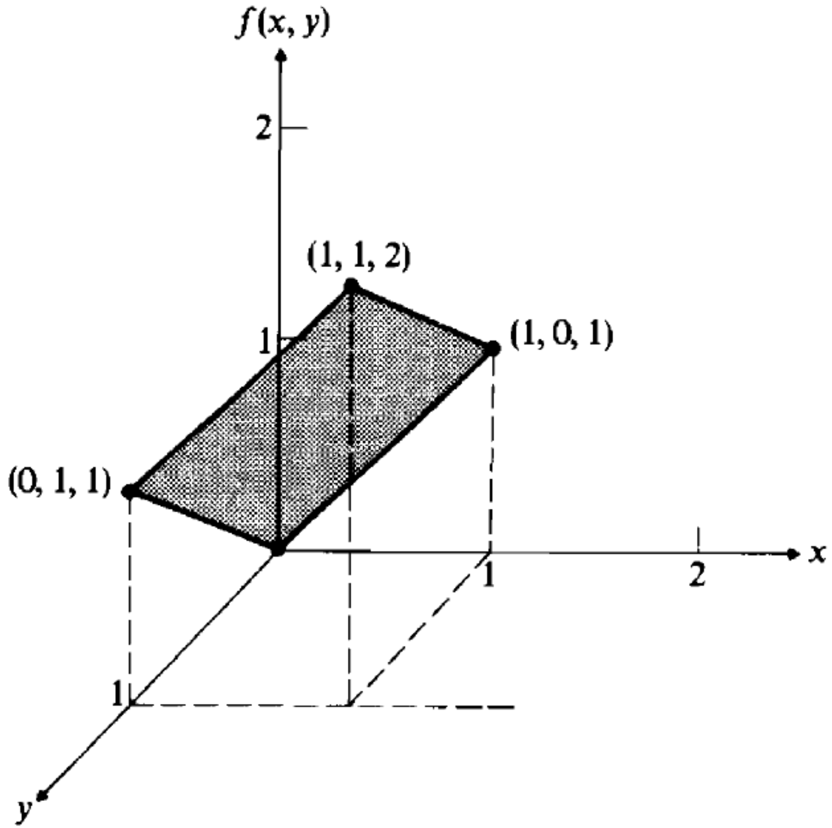
EXEMPLO: Considere a função bivariada dada por

$$f(x,y) = c(x+y)I_{(0,1)}(x)I_{(0,1)}(y) = c(x+y)/I_U(x,y)$$

sendo $U = \{(x,y)|0 < x < 1 \text{ e } 0 < y < 1\}$. Calcule o valor de c para $f(x,y)$ ser considerada uma densidade.

$$\begin{aligned}
1 = \int_0^1\int_0^1 f(x,y)dxdy &= \int_0^1\int_0^1 c(x+y)dxdy \\
&= c\int_0^1\int_0^1 (x+y)dxdy = c\left(\int_0^1 xdx + \int_0^1 ydy\right) \\
&= c\left(\frac{x^2}{2}\Big|_0^1 + \frac{y^2}{2}\Big|_0^1\right) \\
&= c\left(\frac{1}{2} + \frac{1}{2}\right) \implies c = 1
\end{aligned}$$

Função probabilidade conjunta para variáveis aleatórias contínuas



Função probabilidade conjunta para variáveis aleatórias contínuas

Resultado: Se X e Y são variáveis aleatórias contínuas conjuntamente, então o conhecimento de $F_{X,Y}(\cdot,\cdot)$ é equivalente ao conhecimento de $f_{X,Y}(\cdot,\cdot)$. A observação se estende a variáveis aleatórias contínuas k-dimensionais.

$$F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(x,y)dxdy$$

$$f(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$$

Distribuição Normal Multivariada

Distribuição Normal Multivariada

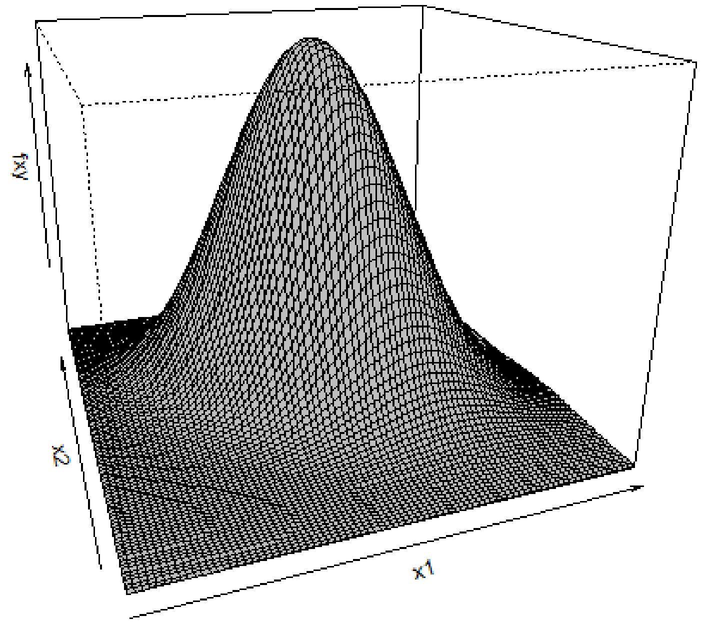
Um vetor aleatório contínuo $\mathbf{X} = (X_1, \ldots, X_k)$ com possíveis valores $\mathbf{x} = (x_1, \ldots, x_k)$ segue uma distribuição normal multivariada com parâmetros $\boldsymbol{\mu} \in \mathbb{R}^{\Bbbk}$ e Σ, $\mathbf{X} \sim N_k(\mu, \Sigma)$ se sua f.d.p multivariada é dada por

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{k}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2} (\mathbf{x} - \mu)' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \mu) \right\}$$

Comentário: $\mathbf{X}$ tem distribuição normal multivariada se qualquer combinação linear de seus componentes $\sum_{j=1}^{p} a_j X_j$ tem uma distribuição normal.

Distribuição Normal Multivariada

Representação gráfica de $k = 2$.



A SHORT REVIEW:

- Distribuição conjunta
- Propriedades da função de distribuição acumulada bivariada
- Função probabilidade conjunta para variáveis aleatórias discretas
- Distribuição multinomial
- Função densidade conjunta para variáveis aleatórias contínuas
- Distribuição normal multivariada